Aalborg Universitet

A multimodal architecture for simulating natural interactive walking in virtual environments

Nordahl, Rolf; Serafin, Stefania; Turchet, Luca; Nilsson, Niels Christian

Published in:
PsychNology Journal

Publication date:
2011

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Nordahl, R., Serafin, S., Turchet, L., & Nilsson, N. C. (2011). A multimodal architecture for simulating natural interactive walking in virtual environments. *PsychNology Journal*, *9*(3), 245-268. http://207.210.83.249/psychnology/File/PNJ9(3)/PSYCHNOLOGY_JOURNAL_9_3_NORDAHL.pdf

PsychNology Journal, 2011
Volume 9, Number 3, 245 – 268

A multimodal architecture for simulating natural interactive walking in virtual environments

Rolf Nordahl[1]*, Stefania Serafin[1], Luca Turchet[1] and Niels C. Nilsson[1]

[1]Dept. of Architecture, Design and Media Technology,
Aalborg University, Copenhagen
(Denmark)

ABSTRACT

We describe a multimodal system that exploits the use of footwear-based interaction in virtual environments. We developed a pair of shoes enhanced with pressure sensors, actuators, and markers. These shoes control a multichannel surround sound system and drive a physically based audio-haptic synthesis engine that simulates the act of walking on different surfaces. We present the system in all its components, and explain its ability to simulate natural interactive walking in virtual environments.

We describe two experiments where the possibilities offered by the system are tested. In the first experiment, blindfolded subjects are asked to walk on a virtual rope, guided only by auditory, haptic and audio-haptic feedback provided at feet level. In the second experiment, subjects are overlooking a virtual canyon, while wearing a head mounted display and the developed shoes. Results of the experiments provide some preliminary indications on the role of multimodal feedback delivered at feet level to enhance realism and sense of presence in virtual environments.

Keywords: *Walking, multimodal interaction, physical models, presence.*

Paper received 02/11/2011; received in revised form 20/12/2011; accepted 22/12/2011.

1. Introduction

During everyday life we routinely navigate the environments we inhabit by walking. For the most part we do so with relative ease and with little or no explicit attention assigned to the movements we perform or the sensory stimuli produced as a result of these movements. However, facilitation of this mundane task is oftentimes anything but a trivial matter in relation to virtual environments. While the use of input devices such as a joystick, mouse or keyboard may facilitate effective interaction, this neither allows

Cite as:
Nordahl, R., Serafin, S., Turchet, L., & Nilsson, N.C. (2011). A multimodal architecture for simulating natural interactive walking in virtual environments. *PsychNology Journal, 9*(3), 245 – 268. Retrieved [month] [day], [year], from www.psychnology.org.

*Corresponding Author :
Stefania Serafin,
Department of Architecture, Design and Media Technology, Aalborg University Copenhagen, Lautrupvang 15, 2750 Ballerup, DK.
E-mail: sts@create.aau.dk

for the transfer of navigational skills acquired during real world walking, nor allows the user to naturally interact. Moreover, the sensation of presence may be positively influenced by the ability to navigate virtual environments in the same way as one would act in real environments. Indeed, it has been shown that the extent to which the locomotion technique resembles its real world correlate has a positive influence on the sensation of presence (Slater, Usoh & Steed, 1995; Usoh et al.1999).

The task of walking in virtual worlds can be broken down into at least two constituent parts that pose two separate, yet interrelated, challenges for creators of walking simulations. First, the real world movement of the user has to result in appropriate virtual egocentric motion. This may pose a problem since the virtual environment's size is usually larger than the users' real world, where movements are confined to a limited physical space. Secondly, the user has to experience appropriate multimodal feedback as a result of the interaction with the virtual environment.

Indeed, when walking in the real world, we receive several kinds of feedback including visual feedback, haptic feedback at the feet that indicates the kind of surfaces we are stumping upon, and auditory feedback connected to our footsteps as well as provided by the environment surrounding us. A virtual walking experience reproducing the real world should be able to simulate all these types of feedback.

Within the academic community several solutions for translating users' real world movement into appropriate virtual egocentric motion have been proposed. Indeed, work pertaining to foot-based interactions has mostly been concerned with the engineering of locomotion interfaces for virtual environments (Pelah & Koenderink 2007). Generally these solutions seem to deal with the physical constraints on the users' movement in one of two ways. They either involve elaborate mechanical setups intended to facilitate natural walking while the user remains at the same physical position, or else they are based on alternative interaction strategies allowing the user to navigate the virtual environment by performing walking-like body movements that does not require actual movement. Examples of the former include omnidirectional treadmills (Darken, Cockayne & Carmein, 1997; Iwata & Yoshida, 1999). Another example is the Virtusphere, that enables users to walk in all directions by placing them inside a large, rotatable, hollow sphere (Medina, Fruland & Weghorst, 2008). The CirculaFloor is an active floor consisting consists of four robotic titles that can reposition themselves thereby allow the user to walk in any direction (Iwata, Yano, Fukushima & Noma, 2005). As a last example, the String-walker combines wheeled shoes with strings actuated by motor-pulley mechanisms in order to facilitate omnidirectional movement

(Iwata, Yano, & Tomiyoshi, 2007). Examples of interaction strategies used to achieve the same goal are the different variations of so-called walking-in-place techniques (e.g., Feasel, Whitton, & Wendt, 2008; Slater, Usoh & Steed, 1995) that enable the user to navigate in virtual environments by walking in place.

The second challenge facing creators of walking simulation is the fact that the user has to experience appropriate multimodal feedback as a result of the interaction with the virtual environment. For example, if we consider audition, when exploring a place by walking, two main categories of sounds can be identified: the person's own footsteps and the surrounding soundscapes. In the movie industry, footstep sounds represent important elements. Chion writes of footstep sounds as being rich in what he refers to as materializing sound indices – those features that can lend concreteness and materiality to what is on-screen, or contrarily, make it seem abstracted and unreal (Chion, Gorbman, & Murch, 1994). We believe that footstep sounds, as well as stimulation of the haptic modality, similarly represent an important element in interactive entertainment, and novel foot-based interactions present new possibilities in this area. With an outset in the writings of Gibson (Gibson, 1986), Slater and colleagues describe bodily movement in terms of the proprioceptive sensory data loop and highlight the importance of this loop in connection to the simulations of a convincing body movement (Slater, Usoh & Steed, 1995). They provide an example of the significance of the loop by describing that when moving a leg so that it touches an object, it is necessary for the individual to receive sensory data, in all modalities, that correspond to the proprioceptive information resulting from the movement. To be more precise, the sensory data is necessary in order to inform the individual that the movement and contact with the object is indeed taking place. While work pertaining to foot-based interactions primarily has been concerned with the engineering of locomotion interfaces, exceptions do exists. However, it would appear that research on the multimodal feedback associated with walking based locomotion interfaces still is in its infancy. Existing interfaces can be categorized as either floor-based or wearable systems.

While not explicitly related to the act of walking in virtual environments, Pinkston (Pinkston, 1994) describes a floor-based solution that transforms user movement into task specific feedback. The system does more specifically function as a touch sensitive dance floor/MIDI controller that captures the user's movements by means of force resisting sensors and transforms these into auditory and visual feedback. Law and colleagues describe a floor-based solution that is able to simulate the experience of walking on different surfaces by means of visual, auditory and haptic feedback (Law, Peck, Visell, Kry, & Cooperstock, 2008). This system consists of a CAVE like environment to which a fourth dynamic multimodal surface has been added, that is, the floor. The user’s movement is tracked by means of a motion capture system. When using this system the user is able to see, hear and feel the surface’s deformations produced when he or she is walking within the environment.

In regards to wearable solutions, Paradiso and coworkers pioneered the development of shoes enhanced with sensors (Paradiso, Hsiao & Hu, 1999). The developed shoes were able to capture 16 different parameters such as pressure, orientation, and acceleration and were intended for musical performances as well as for rehabilitation studies (Benbasat, Morris & Paradiso, 2003). Notably, the company Nike has also developed the Nike+ sensor (http://nikerunning.nike.com), which is an accelerometer that can be attached to one’s running shoes and connected wirelessly to an iPod or iPhone. The sensor is then able to provide the user with relevant information about the running activity via the iPod or iPhone.

In this paper we describe a multimodal interactive space relying on a wearable solution. This space has been developed with the intention of creating audio-haptic-visual simulations of walking-based interactions. Compared to previous solutions, this system presents for the first time a multimodal environment where both auditory and haptic feedback are delivered using physics based modeling, and are complemented by visual feedback. Moreover, all three kinds of modalities – audition, vision and touch, are present both as input and output. The system requires users to walk around a space wearing a pair of shoes enhanced with sensors and actuators. The position of such shoes is tracked by a motion capture system, and the shoes drive an audio-visual-haptic synthesis engine based on physical models. An interesting feature of this system is that it allows for relatively easy integration with most of the described locomotion interfaces.

We have used this architecture to perform several psychophysical experiments in order to understand the contribution of the auditory and haptic modalities when interacting with different simulated surfaces using the feet (Turchet et al. 2010c). We have also investigated the role of the different modalities when providing feedback in balancing tasks, as well as the possibility of recreating sense of presence in virtual environments (Nordahl, 2010).

Possible applications of the architecture are envisaged in the field of navigation in real and virtual environments, architecture, rehabilitation and entertainment. As an example, a better understanding of the role of the different modalities in helping balance control can advance the field of virtual reality for rehabilitation purpose. Moreover, the possibility to have faithful reproduction of real places, both indoors and outdoors, is an advance in the field of virtual reality for architecture, as well as the ability to visit a physical place virtually. Additionally, in the entertainment industry, several interfaces such as the Wii Fit by Nintendo (www.wiifit.com) and the Kinect by Microsoft (www.xbox.com/kinect) are starting to explore the possibilities offered by feet-based and full-body interactions. Amusement parks are also exploring the possibilities offered by virtual reality and multimodal interaction in order to provide illusions such as vection, i.e., the illusion of self-movement in space.

The paper is organized as follows: Section 2 describes the hardware of the developed architecture, and Section 3 its software. Section 4 presents two experiments where the architecture has been adopted. These experiments investigate the role of multimodal feedback in feet-based interactions, and specifically whether haptic feedback enables improvements in performance, perceived realism and sense of presence. Section 5 presents the conclusions.

2. The overall architecture

The main goal of the developed architecture is to create a multimodal input-output system able to track the position of the users' shoes and head in order to drive

An audio-haptic synthesis engine based on physical models and supported by visual feedback. In order to achieve this goal, we chose a motion capture system to track the users' feet, and we developed some custom made shoes able to provide haptic feedback, as described later.

The architecture consists of a motion capture system (MoCap), two soundcards, twelve loudspeakers, two amplifiers, two haptic shoes, a head-mounted display (HMD), and two computers. The system is placed in an acoustically isolated laboratory that consists of a control room and a bigger room where the setup is installed and where the experiments are performed. The control room is 5.45 m large, 2 m long, and 2.85 m high, and it is used by the experimenters providing the stimuli and collecting the experimental results. It hosts two desktop computers. The first computer runs the motion capture software (Tracking Tools 2.0) and the visual engine Unity 3D, while the second runs the audio-haptic synthesis engine. The two computers are connected through an Ethernet cable and communicate by means of the UDP protocol (http://opensoundcontrol.org/). The data relative to the motion capture system are sent from the first to the second computer that processes them in order to control the sound engine. A transparent glass divides the two rooms, so it is possible for the experimenters to see the users performing the assigned task. The two rooms are connected by means of a standard talk- back system such as the ones used in recording studios. The experiment room is 5.45 m large, 5.55 m long, and 2.85 m high, and the walking area available to the users is about 24m^2.

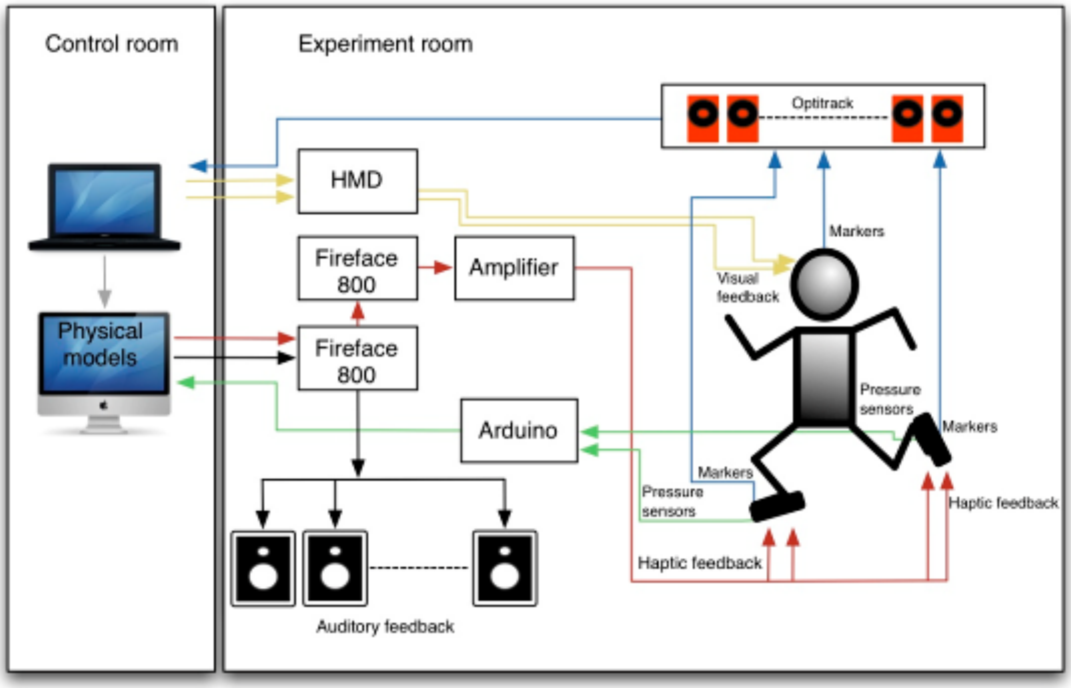

Figure 1. A schematic representation of the developed multimodal architecture

2.1 Tracking the User

The user's locomotion is tracked by an Optitrack motion capture system (http://naturalpoint.com/optitrack/) composed by 16 infrared cameras (OptiTrack FLEX: V100R2). The cameras are placed in a configuration optimized for the tracking of the feet and head position simultaneously. Following recommendation from the cameras' manufacturers, we placed eight cameras close to the ceiling, pointed towards the center of the room to track the participants head, and eight cameras close to the floor, pointing to the center of the room to track the participants' feet. Eight cameras were attached to the eight vertexes of a square metal frame. The other eight cameras were attached in between the other cameras in the horizontal frames.

In order to achieve the goal of tracking the feet, markers were placed on the top of each shoe worn by the users as well as on top of the head. Users are also tracked by using the pressure sensors embedded in a pair of sandals, as shown in Figure 2 and Figure 3. Specifically, a pair of light- weight sandals was used (Model Arpenaz-50, Decathlon, Villeneuve d'Ascq, France).

The sole has two FSR pressure sensors (I.E.E. SS-U-N-S- 00039) whose aim is to detect the pressure force of the feet during the locomotion of a participant wearing the shoes. The two sensors are placed in correspondence to the heel and toe respectively in each shoe. The analogue values of each of these sensors are digitalized by means of an Arduino Diecimila board (http://arduino.cc/) and are used to drive the audio and haptic synthesis.

2.2 Haptic Feedback

In order to provide haptic feedback during the act of walking, a pair of custom made shoes with sensors and actuators has been recently developed (Turchet et al., 2010a). The particular model of shoes chosen has light, stiff foam soles that are easy to gouge and fashion. Four cavities were made in the thickness of the sole to accommodate four vibrotactile actuators (Haptuator, Tactile Labs Inc., Deux-Montagnes,Qc, Canada). These electromagnetic recoil-type actuators have an operational, linear bandwidth of 50–500 Hz and can provide up to 3G of acceleration when connected to light loads.

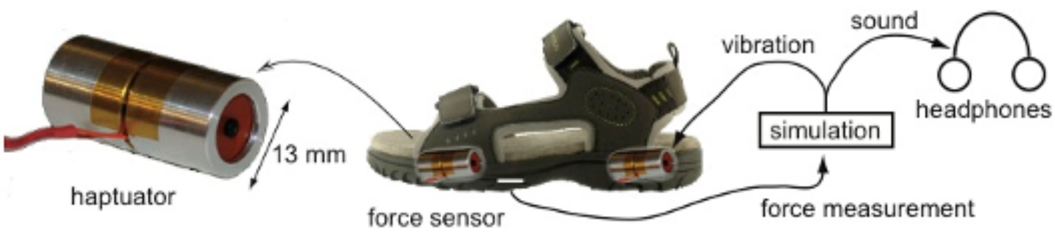

Figure 2. The developed haptic shoes used as part of the multimodal architecture

Figure 3. A picture of one pressure sensor and two actuators embedded in the shoes.

As indicated in Figure 2 and Figure 3, two actuators were placed under the heel of the wearer and the other two under the ball of the foot. These are bonded in place to ensure good transmission of the vibrations inside the soles. When activated, vibrations propagated far in the light, stiff foam. In the present configuration, the four actuators are driven by the same signal but could be activated separately to emphasize, for instance, the front or back activation, or to realize other effects such as modulating different back-front signals during heel-toe movements. A cable exits from each shoe, with the function of transporting the signals of the pressure sensors and for the actuators. These cables are about 5 meters long, and they are connected, through DB9 connectors, to two 4TP (twisted pair) cables: one 4TP cable carries the sensor signals to a breakout board, which then interfaces to an Arduino board; the other 4TP cable carries the actuator signals from a pair of Pyle Pro PCA1 mini 2X15 W stereo amplifiers, driven by outputs from a FireFace 800 soundcard. Each stereo amplifier handles 4 actuators found on a single shoe, each output channel of the amplifier driving two actuators connected in parallel. The PC handles the Arduino through a USB connection, and the FireFace soundcard through a FireWire connection.

2.3 Auditory Feedback

In our virtual environment auditory feedback can be delivered by means of headphones or a set of loudspeakers (Dynaudio BM5A speakers).

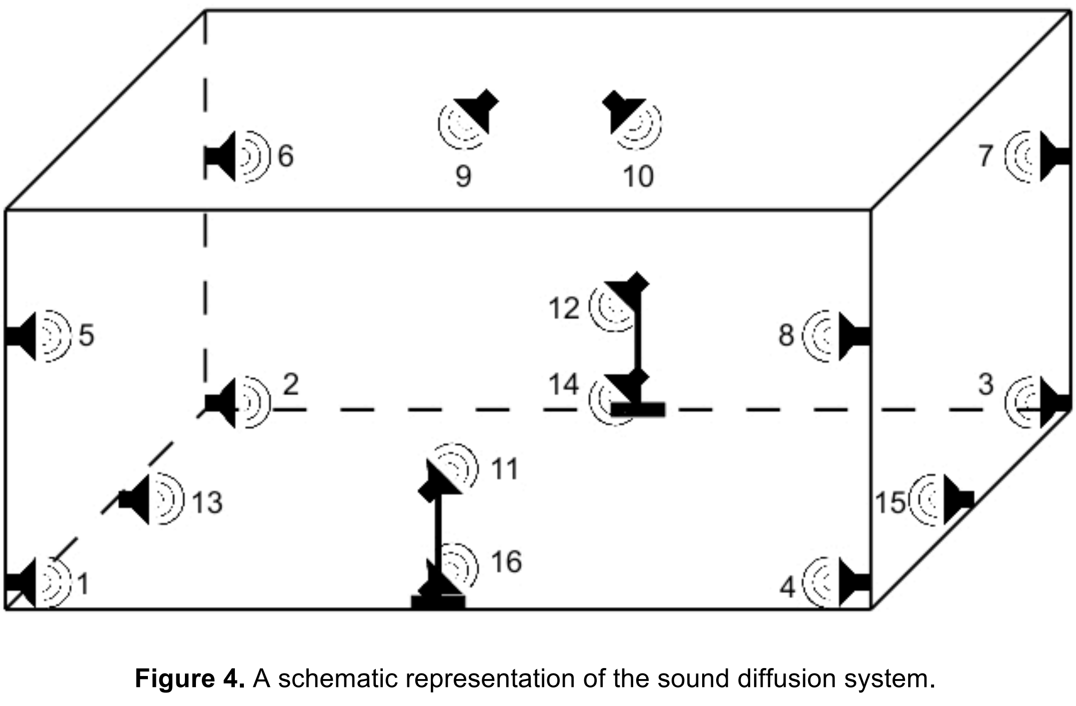

Figure 4. A schematic representation of the sound diffusion system.

The loudspeakers' configuration is illustrated in Figure 4. This configuration was chosen as one of the several possible solutions in order to acoustically cover the different sides of the room. In the current setup we use 12 or 16 loudspeakers depending on whether the haptic feedback is involved or not. Indeed for the delivery of both the haptic and auditory feedback we use two FireFace 800 soundcard connected through a firewire 800 cable (see Figure 1). Since there are 8 output channels available on each soundcard, and handling the haptic feedback requires four output channels, we use the remaining 12 for the auditory feedback (loudspeakers 1-12 in Figure 5). Conversely, when the haptic feedback is not involved all the 16 channels are available for the auditory feedback (loudspeakers 1-16 in Figure 5). In the future, we plan to extend this configuration adding a third soundcard dedicated exclusively to the handling of the haptic feedback.

2.4 Visual Feedback

The visual feedback is provided by a head-mounted display (HMD) nVisor SX from nVis (www.nvis.com). The HMD is connected to the PC by using a Matrox TripleHead2Go Digital Edition graphics expansion module. As previously mentioned, three markers are placed on top of the HMD, in order to track orientation and position of the head. The goal of the visual feedback is to render, through the use of a commercial game engine, the visual sensation of exploring different landscapes. In particular, in our simulation the Unity3D game engine has been used (http://unity3d.com/). This engine was used because of its ability to render realistic visual environments without being skilled visual designers. This choice was ideal for us, since our main interest is a physically based audio-haptic engine, so the visual feedback is used only for supporting it, without being the main goal. Simple dynamic stereoscopy was implemented. Eye convergence was simulated by using a raycasting algorithm, which ensures that the cameras are always aimed at the closest object immediately in front of the user. This choice was ideal for us, since our main interest is a physically based audio-haptic engine, so the visual feedback is used only for supporting it, without being the main goal.

3. Simulation Software

We developed a multimodal synthesis engine able to reproduce auditory and haptic feedback. As concerns the auditory feedback, we developed a sound synthesis engine based on a footstep sounds synthesis engine and on a soundscapes synthesis engine.

Figure 5. A screenshot of the audio-haptic synthesis engine used in the architecture.

The engine for footstep sounds, based on physical models, is able to render the sounds of footsteps both on solid and aggregate surfaces. Several different materials have been simulated, in particular wood, creaking wood, and metal as concerns the solid surfaces, and gravel, snow, sand, dirt, forest underbrush, dry leaves, and high grass as regards the aggregate surfaces. A complete description of such engine in terms of sound design, implementation and control systems is presented in (Turchet, Serafin, Dimitrov & Nordahl 2010c).

Using this engine, we implemented a comprehensive collection of footstep sounds. As solid surfaces, we implemented metal, wood, and creaking wood. In these materials, the impact model was used to simulate the act of walking, while the friction model was used to simulate the creaking sounds typical of creaking wood floors. As aggregate surfaces, we implemented gravel, sand, snow, forest underbrush, dry leaves, pebbles and high grass. The simulated metal, wood and creaking wood surfaces were furthermore enhanced by using some reverberation. To control the audio-haptic footsteps synthesizer in our virtual environment, we use the haptic shoes: the audio-haptic synthesis is driven interactively during the locomotion of the participant wearing the shoes. The description of the control algorithms based on the analysis of the values of the pressure sensors coming from the haptic shoes can be found (Turchet et al.2010a). This engine has been extensively tested by means of several audio and audio-haptic experiments and results can be found in Nordahl, Serafin and Turchet (2010b), Nordahl et al. (2010a), Serafin et al. (2010) and Turchet, Nordahl and Serafin (2010b). The graphical user interface of the sound synthesis engine can be seen in Figure 5.

3.1 Soundscape Rendering

In order to sonically render the sensation of walking in different locations, we implemented a soundscape engine able to provide various typologies of soundscapes: static soundscapes, dynamic soundscapes and interactive soundscapes. Static soundscapes are those composed without rendering the appropriate spatial position of the sound sources, nor their tridimensional movements in the space. An example of a static soundscape is a soundscape where each speaker delivers the same sounds at the same amplitude, no matter where the user is placed. Conversely, in the dynamic soundscapes the spatial position of each sound source is taken into account, as well as their eventual movements along tridimensional trajectories. Finally, the interactive soundscapes are based on the dynamic ones where in addition the user can interact with the simulated environment generating an auditory feedback as result of his/her actions. An example of interactive soundscape is a soundscape where when a user walks in the physical environment, a footstep sounds is heard, together with the environmental sounds of the simulated place. The position and the movements of the user are tracked by means of the MoCap system and are used as input for the designed interactive soundscapes. As an example of sound interaction, one can imagine the situation in which the virtual environment simulates a forest, and when the user walks close enough to a bush where there are some animals the sounds of the movements of the animals are triggered. Furthermore, the footstep sounds interactively generated during the locomotion of the user can be conveyed to the user taking into account the position of the user's feet in the simulated space, in such a way that the footstep sounds the user's position. In this way, the user can perceive as if the footstep sounds are coming directly from their source, instead of coming from the speakers.

The sound synthesis engine has been implemented using the Max/MSP software platform by Cycling 74. To achieve the dynamism in the soundscapes we use the ambisonic tools for Max/MSP, which makes it possible to move virtual sound sources along trajectories defined on a tridimensional space (Schacher & Neukom, 2006). In particular, the sound synthesis engine is currently set with 16 independent virtual sound sources, one for the footstep sound, and fifteen for the sound sources present in the soundscape.

3.2 Haptic Feedback

Concerning the haptic feedback, it is provided by means of the haptic shoes previously described. The haptic synthesis is driven by the same engine used for the synthesis of footstep sounds, and is able to simulate the haptic sensation of walking on different surfaces, as illustrated in (Turchet et al., 2010a).

3.3 Visual Feedback

As regards to the visual feedback, several outdoor scenarios have been developed using the Unity3D engine. The goal of such outdoor scenarios is to provide a visual representation of the physically simulated surfaces provided in the audio-haptic engine. As an example, a forest, a beach and a ski slope were visually rendered, to match the physically simulated sand, forest underbrush and snow. The user interacts with the visual engine by means of the markers placed on the top of the HMD, and by means of the pressure sensors embedded in the shoes.

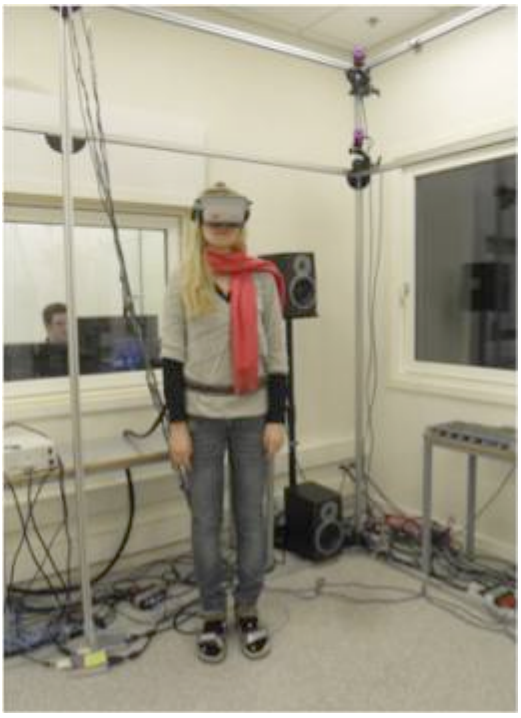

Figure 6. A participant interacting with the developed architecture

Figure 6 shows a participant interacting with the virtual environment. The participant is wearing the HMD and the shoes enhanced with actuators, pressure sensors and markers. In the background it is possible to notice the motion capture cameras as well as the surround sound system.

4.Experiment Design

We designed two experiments that evaluate multimodal feedback delivered at feet level. The first experiment does not present any visual feedback and requires subjects to be blindfolded. In the second experiment, visual feedback is also present and delivered through the HMD previously described.

4.1 Experiment 1: Walking on a Virtual Rope

The goal of this experiment is to understand whether auditory and haptic feedback facilitates the task of walking on a virtual rope.

4.1.1 Procedure

Participants were asked to wear the haptic sandals previously described and to walk blindfolded straight in order not to fall from a virtual plank. Figure 7 shows a participant performing the experiment. Specifically, participants were given the following instructions: "Imagine you are walking on a wooden plank. Your task is to walk from one side to the other. Walk slowly and pay attention to the feedback you receive in order to succeed on your task. If your feet are outside of the plank you will fall." The same stimuli were provided for the auditory and haptic simulation and designed as follows: when a user was walking on top of the virtual plank, his position was detected by the motion capture system previously described. In this case, the synthesis engine provided as a stimulus the sound and haptic feedback of a creaking wood. The physics based synthesis engine was implemented using the algorithms described in (Nordahl, Serafin & Turchet, 2010b).

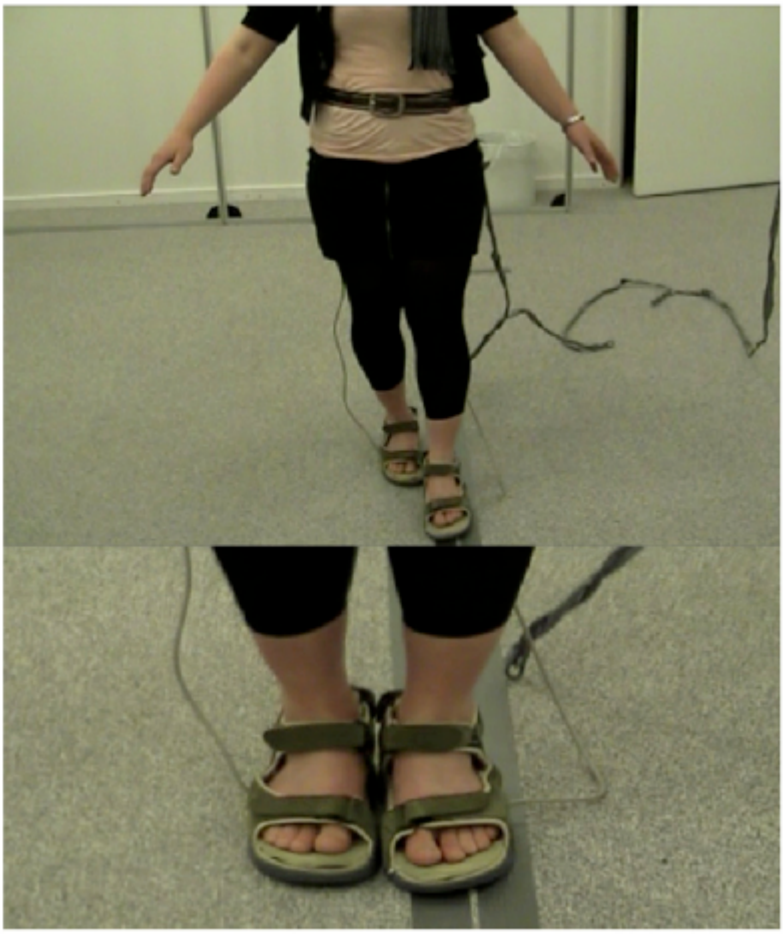

Figure 7. A participant performing the experiment consisting of walking on a virtual plank

4.1.2 Participants

The experiment was performed by 15 participants, 14 men and 1 woman, aged between 22 and 28 (mean=23.8, st.d.=1.97). All participants reported normal hearing conditions. The experiment was conducted as a within-subjects experiment, where subjects were randomly exposed to the four following conditions: auditory feedback, haptic feedback, audio-haptic feedback and no feedback. Each condition was experienced twice, giving in total eight trials for each subject.

4.1.3 Results and Discussion

Table 1 shows the performance for each participant. The numbers in each row for each condition indicate whether the participant performed successfully the task ones, twice or never.

Condition/ Subject number	Audio (A)	Haptic (H)	Audio-haptic (AH)	No-feedback (N)
1	2	1	2	1
2	2	1	1	1
3		1		
4	1	1		
5	1		2	
6	1	1	2	1
7		2	1	
8	1			
9	1	1		
10				
11	1	1		
12		2	1	1
13				
14		1	2	1
15	2	2	1	2

Table 1. Summary of the results of the experiment consisting on walking on a virtual rope. The number in each element of the matrix represents the times the task was successful (once, twice or never).

The results show that feedback helps balance mostly when haptic stimuli are provided. In this case, 46.6% of the tasks were successfully completed. In the case where a combination of auditory and haptic feedback was provided, 43.3% of the tasks were completed. With only auditory feedback, 40% of the tasks were completed, while with no feedback only 26.6%. These results show that feedback slightly helps the balancing task. Haptic feedback performed better than the combination of auditory and haptic. This can be due to the fact that haptic feedback was provided directly at the feet level, so the participants had a closer spatial connection between the path they had to step on and the corresponding feedback.

A post-experimental questionnaire was also performed, where participants were asked several questions regarding their ability to freely move in the environment, to adjust to the technology and to which feedback was the most helpful. Indeed, 7 participants found the haptic feedback to be the most helpful, 6 participants the auditory feedback and 2 participants the combination of auditory and haptic feedback. One participant commented that the most useful feedback was when there was background noise (the pink noise used to mask the auditory feedback) and only vibration was provided. All participants claimed to notice the relationship between actions performed and feedback provided.

Some participants also commented on the fact that shoes were not fitting their size. Moreover, some felt disable without the visual feedback. One participant observed that he simply ignored the feedback and walked straight. This is an indication of his unwillingness of suspending his disbelief, and behave in a way similar to how they would behave when walking on a real narrow plank [15].

Overall, observations of most of the participants showed that they were walking carefully listening and feeling the feedback in order to successfully complete the task. It is hard to assess whether the lack of feedback was the condition participants were exposed to, the fact that they were outside the plank or a fault of the system. Some of the test participants were noticeably not walking straight, although in the post-experimental questionnaire they commented on a faulty system. Very few understood that the lack of feedback was provided intentionally.

4.2 Experiment 2: Enhancing Presence and Realism through Audio-Haptic Feedback

We designed an experiment whose goal was to investigate the role of auditory and haptic feedback in enhancing presence and realism in a virtual environment. As can be seen in Figure 8, in the first conditions participants were asked to stand on a physical wooden plank while experiencing the environment. The same plank was not present in the second condition. The reason was to investigate whether passive haptic, defined as the augmentation of virtual environments with low-fidelity physical objects, had an effect in the results. The visual feedback the participants were exposed to is displayed in Figure 9. In order to allow participants to explore the environment to its entirety, and ensure that they approached the edge of the platform while looking down, participants were asked to look and find three objects located underneath the virtual platform.

In each condition, half of the participants experienced the lack of audio-haptic feedback first and the presence of audio-haptic feedback afterwards, while the other half experienced the presence of audio-haptic feedback first and the lack of audio-haptic feedback afterwards. Audio-haptic feedback was provided using the shoes previously described.

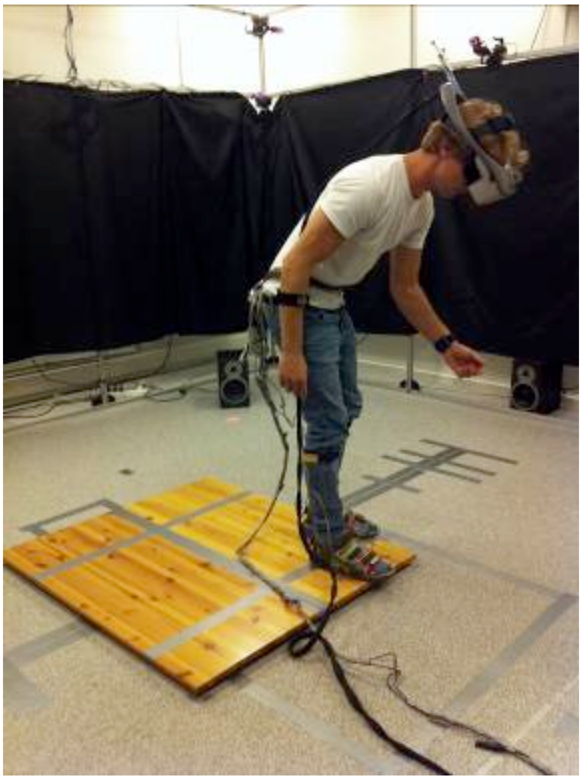

Figure 8. A participant performing the experiment of overlooking the virtual canyon.

Figure 9. A view of the visual feedback provided to the users, where the users' own feet are visible.

4.2.1 Participants

Forty participants were divided in two groups (n=20) to perform the experiment. The two groups were composed respectively of 15 men and 5 women, aged between 20 and 34 (mean=23.05, standard deviation=3.13), and of 15 men and 5 women, aged between 20 and 32 (mean=23.5, standard deviation=3.17). Participants were primarily recruited from the campus of the Media Technology Department of Aalborg University Copenhagen; however no restrictions on background were imposed. All participants reported normal, or corrected to normal, hearing.

4.2.2 Results

Participants' behavior was measured by recording participants' heart rate, galvanic skin response and skin conductance. The participants' experience of presence was assessed by means of the Slater-Usoh-Steed (SUS) questionnaire (Usoh, 2000). In this paper, we report only the results gathered through the presence questionnaire. This questionnaire is intended to evaluate the experience after exposures to a virtual environment (VE). The SUS questionnaire contains six items that evaluate the experience of presence in terms of, the participants' sense of being in the VE, the extent to which the participant experienced the VE as the dominant reality, and the extent to which the VE is remembered as a place. All items are answered on scales ranging from 1 to 7 where the highest scores would be indicative of presence (Usoh, 2000):

Q1: Please rate your sense of being in the virtual environment, on a scale of 1 to 7, where 7 represents your normal experience of being in a place.

Q2: To what extent were there times during the experience when the virtual environment was the reality for you?

Q3: When you think back to the experience, do you think of the virtual environment more as images that you saw or more as some- where that you visited?

Q4: During the time of the experience, which was the strongest on the whole, your sense of being in the virtual environment or of being elsewhere?

Q5: Consider your memory of being in the virtual environment. How similar in terms of the structure of the memory is this to the structure of the memory of other places you have been today?

Q6: During the time of your experience, did you often think to your- self that you were actually in the virtual environment?

Moreover, during the experiment skin conductance, skin temperature and heart rate were measured.

The general level of presence experienced by the participants may be determined by summarizing the data obtained from all of the questionnaire items in two ways. First, one may present the central tendency as the mean of all ratings to all items and the variability may thus be presented as the corresponding standard deviation. Secondly, it is possible to present the general experience of presence across participants (SUS count), as the mean of the individual presence scores. The presence score is taken as the sum of scores of 6 and 7 out of the number of questions posed. Tables 2 and 3 illustrate the questionnaire evaluations for the first and second condition respectively. In the tables, NF indicates the trial with no feedback, while F indicates the trial with feedback.

	Trials NF-F		Trials F-NF	
	NF	F	NF	F
Q1	5.3±1.49	6±1.15	5.63±1.2	5.45±1.12
Q2	5.5±1.08	5.6±1.26	5.45±1.36	5.45±1.12
Q3	3.9±1.79	5.1±1.59	5.09±1.57	5.81±0.98
Q4	5.9±0.99	6.1±0.56	5.18±1.53	6.18±0.87
Q5	3.5±1.5	4.7±1.63	4.72±1.79	4.54±0.93
Q6	4.9±1.72	5.3±1.7	5.09±2.02	6.18±1.16
SUS count	0.38±2.13	0.65±2.16	0.6±0.89	0.6±2.52

Table 2. Questionnaire's results of the condition with passive haptics.

	Trials NF-F		Trials F-NF	
	NF	F	NF	F
Q1	4.7±1.15	5.4±1.26	5.3±1.56	5.6±0.96
Q2	4.±1.15	5.±1.33	4.8±1.22	5.3±0.94
Q3	4.5±1.77	4.7±1.56	4.8±1.68	4.8±1.68
Q4	5.4±1.07	5.3±1.15	5.5±0.97	5.3±1.15
Q5	3.9±2.55	4.3±2.49	4.1±1.79	4.2±2.14
Q6	3.6±1.57	5.2±1.68	5.7±1.25	4.9±1.37
SUS count	0.28±1.17	0.5±1.41	0.46±1.75	0.41±0.75

Table 3. Questionnaire's results of the condition without passive haptics.

As outlined in (Usoh, 2000), to check if the differences found in the questionnaire results for the two typologies of stimuli F and NF are statistically significant, one should not compare the means of the questionnaire's items results, but rather the number of answers having a score of 6 or 7. Following this approach we found statistical significance in both conditions (with and without passive haptics) for the trials in which the no feedback condition was presented first and the feedback condition afterwards ($\chi^2(1) = 5.0364$, p-value = 0.02482 and $\chi^2(1) = 7.5083$, p-value = 0.006141 respectively). Conversely, no significance was found in any of the two conditions for the trials in which the feedback condition was presented first and the no feedback condition afterwards. It is interesting to notice that the mean presence score pertaining to the feedback condition is significantly higher when this condition was presented first while there was no significant difference between the scores when the no-feedback condition was presented first, despite this average being the higher. One may argue that this lends some credence to the claim that the addition of the feedback did increase the participants' sensation of presence. It does therefore remain an open question whether the added feedback did in fact increase the sensation of presence on behalf of the participants. With this being said, it is worth noting that results obtained from the questionnaire at least in part correspond with the statements made by the participants who generally thought that the feedback added to the sense of realism and in some cases intensified the experience of vertigo. Moreover, while the choice of the SUS-presence questionnaire was motivated by the fact that it is extensively validated and used in the VR community, it can be questioned whether it is the most suitable for examining the relationship between feedback and presence.

As a final analysis of the experiments' results, it is interesting to discuss the observations provided by the participants when the experiments were completed. Specifically, we asked participants if they had noticed any difference on the two conditions and, in affirmative case, if they could elaborate on the differences noticed and how they affected their experience. During the first experiment, when asked whether they had noticed a difference between the two trials, 13 of the participants mentioned that they had noticed the change in the haptic and/or auditory feedback provided by the shoes. Precisely, 5 participants noticed a difference in both auditory and haptic feedback, 7 only noticed the difference in auditory feedback, while 1 only noticed the difference in haptic feedback. All of the participants who noticed the difference expressed a preference towards the added feedback. When asked to elaborate, 11 of the 13 stated that it added realism, 5 felt that it made the experience scarier or intensified the sensation of vertigo, while 1 explicitly stated that it increased the sensation of presence in the virtual environment.

During the second experiment, out of the 20 participants, 16 noticed the additional feedback, 5 participants noticed both the auditory and haptic feedback while 7 just noticed the sound and 4 only noticed the haptic feedback. With one exception, all of the participants who noticed the difference preferred the additional feedback. The one participant who did not, described that he did like the haptic feedback, but he had found it too intense. Out of the 16 who noticed the feedback 13 thought that it added realism, 2 described that it made it more scary and 2 explicitly stated that it intensified the sensation of being there. These observations show that participants indeed were able to notice and appreciate the provided feedback in both experimental conditions. The lack of the same evidence while analyzing physiological data or presence questionnaire can be due to the fact that the provided feedback does not necessarily elicit a higher physiological response or sense of presence.

5. Conclusions

In this paper, we have described the different components of a multimodal interactive space driven by walking. Two experiments that exploit the possibilities offered by the architecture have been presented: in the first experiment, participants were blindfolded and asked to walk on a virtual plank driven by auditory and haptic feedback. In the second experiment, participants were exposed to visual feedback of a canyon, and their physiological reaction was measured and the sense of presence evaluated via a post-experimental questionnaire.

While none of the described experiments provide strong indications on the role of different kinds of feedback in facilitating task performance and enhancing sense of presence, none less participants feedback and gathered data support our hypotheses that haptic feedback provided at feet level is appreciated by the participants and enhances perceived realism. We are currently investigating applications of our architecture in the fields of rehabilitation of lower body parts, virtual exploration of real places and entertainment.

6. Acknowledgments

The results presented in this paper are part of the Natural Interactive Walking (NIW) FET – Open EU project (FP7-ICT-222107), whose goal is to provide closed-loop interaction paradigms enabling the transfer of skills that have been previously learned in everyday tasks associated to walking. In the NIW project, several walking scenarios are simulated in a multimodal context, where especially audition and haptic play an important role.

7. References

Benbasat, A., Morris, S., & Paradiso, J. (2003). A wireless modular sensor architecture and its application in on-shoe gait analysis. *Proceedings of the IEEE International Conference on Sensors* volume 2 (pp. 92-120). IEEE Press.

Chion, M., Gorbman, C., & Murch, W. (1994). *Audio-vision: sound on screen*. New York: Columbia Univ Press.

Darken, R.P., Cockayne, W.R., & Carmein, D. (1997). The Omni-Directional Treadmill: A Locomotion Device for Virtual Worlds. *Proceedings of the 10th annual ACM Symposium on User Interface Software and Technology* (pp. 213-221). New York: ACM Press.

Feasel, J., Whitton, M. C., & Wendt, J. D. (2008). LLCM-WIP: Low-Latency, Continuous-Motion Walking-in-Place. *Proceedings of the IEEE Symposium on 3D User Interfaces* (pp. 97-104). IEEE Press.

Gibson, J. J. (1986). *The Ecological Approach to Visual Perception*. Hillsdale, NJ: Lawrence Erlbaum.

Iwata, H., & Yoshida, Y. (1999). Path Reproduction Tests Using a Torus Treadmill. *Presence: Teleoperators and Virtual Environments, 8*(6): 587-597.

Iwata, H., Yano, H., Fukushima, H., & Noma, H. (2005). CirculaFloor. *Computer Graphics and Applications*, *25*(1): 64-67.

Iwata, H, Yano, H., & Tomiyoshi, M. (2007). String walker. *Proceedings of ACM SIGGRAPH 2007 emerging technologies, SIGGRAPH '07* (Article 20). New York: ACM Press.

Law, A. V., Peck, B. V., Visell, Y., Kry, P. J., & Cooperstock J. R. (2008). A Multi-modal Floor-space for Experiencing Material Deformation Underfoot in Virtual Reality. *Proceedings of the IEEE International Workshop on Haptic Audio Visual Environments and Games*, HAVE'08 (pp. 126-132). IEEE Press.

Medina, E., Fruland, R. & Weghorst, S. (2008). Virtusphere: Walking in a Human Size VR "Hamster Ball". *Proceedings of the Human Factors and Ergonomics Society*, *52*(27): 2102-2106.

Nordahl, R. (2010). Evaluating Environmental Sounds from a Presence Perspective for Virtual Reality Applications. *EURASIP Journal on Audio, Speech, and Music Processing* (Article ID 426937, 10 pages).

Nordahl, R., Berrezag, A., Dimitrov, S., Turchet, L., Hayward, V., & Serafin, S. (2010a). Preliminary experiment combining virtual reality haptic shoes and audio synthesis.

In A. M. L. Kappers, J. B. F. van Erp, W. M. B. Tiest, & F. C. T. van der Helm (Eds.). *Proceedings of the 2010 international conference on Haptics - generating and perceiving tangible sensations: Part II (EuroHaptics'10)* (pp. 123-129). Berlin: Springer.

Nordahl, R., Serafin, S., & Turchet, L. (2010b). Sound synthesis and evaluation of interactive footsteps for virtual reality applications. *Proceedings of IEEE Virtual Reality Conference* (pp. 147-153). IEEE Press.

Paradiso, J., Hsiao, K., & Hu, E. (1999). Interactive music for instrumented dancing shoes. *Proceedings of the 1999 International Computer Music Conference* (pp. 453–456). IEEE Press.

Pelah, A. & Koenderink, J. (2007). Editorial: Walking in real and virtual environments. *ACM Transactions on Applied Perception*, *4*(1) article 1.

Pinkston, R. (1994). A Touch Sensitive Dance floor/MIDI Controller. *The Journal of the Acoustical Society of America*, *96*(5), 3302.

Schacher, J. & Neukom, M. (2006). Ambisonics Spatialization Tools for Max/MSP. *Proceedings of the International Computer Music Conference* (pp. 274-277). IEEE Press.

Serafin, S., Turchet, L., Nordahl, R., Dimitrov, S., Berrezag, A., & Hayward, V. (2010). Identification of virtual grounds using virtual reality haptic shoes and sound synthesis. *Proceedings of Eurohaptics symposium on Haptics and Audio-visual environments* (pp. 61-70). IEEE Press.

Slater, M., Usoh, M. & Steed, A. (1995). Taking Steps: The Influence of a Walking Technique on Presence in Virtual Reality. *ACM Transactions on Computer-Human Interaction (TOCHI),* 2(3): 201-219.

Turchet, L., Nordahl, R., Berrezag, A., Dimitrov, S., Hayward, V., & Serafin, S. (2010a). Audio-haptic physically based simulation of walking sounds. *Proceedings of IEEE International Workshop on Multimedia Signal.Processing* (pp. 269 – 273). IEEE Press.

Turchet, L., Nordahl, R., & Serafin, S. (2010b). Examining the role of context in the recognition of walking sounds. *Proceedings of Sound and Music Computing Conference* (6 pages). IEEE Press.

Turchet, L., Serafin, S., Dimitrov, S., & Nordahl, R. (2010c). Physically based sound synthesis and control of footsteps sounds. *Proceedings of Digital Audio Effects Conference* (pp. 161-168).

Usoh, M., Arthur, K., Whitton, M. C., Bastos, R., Steed, A., Slater, M. & Brooks,Jr., F.

P. (1999). Walking > walking-in-place > flying, in virtual environments. *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques,* SIGGRAPH 99 (pp. 359–364). New York: ACM Press.

Usoh, M., Catena, E., Arman, S., & Slater, M. (2000). Presence Questionnaires in Reality.*Presence: Teleoperators and Virtual Environments, 9*(5): 497-503.